Asian Journal of Environment & Ecology

Volume 22, Issue 4, Page 115-119, 2023; Article no.AJEE.110125
ISSN: 2456-690X

Machine Learning Techniques of Weather Forecasting – A Review

Jitiprova Ghosh [a*] and Amitlal Bhattacharya [b]

[a] Assistant Teacher of Physical Science, Aurangabad Balika Vidyalaya (H.S), India.
[b] Department of Physics, Dukhulal Nibaran Chandra College, Aurangabad, Murshidabad, India.

Authors' contributions

This work was carried out in collaboration between both authors. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJEE/2023/v22i4513

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/110125

Review Article

Received: 05/10/2023
Accepted: 08/12/2023
Published: 18/12/2023

ABSTRACT

Weather is a particular state of the atmosphere that describes the degrees to which it is hot or cold, wet or dry, calm or stormy, clear or cloudy. On earth, most weather phenomena occurs in the lowest layer of the planet's atmosphere i.e the troposphere. Weather forecasting tools are used in the field of science and technology to forecast atmospheric conditions for a certain place and period. It is a very challenging task for the researchers of this field in this modern era. In this review paper we have tried to estimate the accuracy level of different machine learning method of weather forecasting.

Keywords: Random forest; decision tree; support vector machine; naïve bays; logistic regression; K-nearest neighbors; weather data.

1. INTRODUCTION

Now a days weather and climate of this planet is changing very drastically and continuously as a result of the climate change [1,2]. Earth is suffering very badly due to this continuous climate change. It is very crucial and essential task to predict the exact current weather conditions. Presently machine learning techniques are widely use to forecast the weather [3-5]. These techniques depend on complex physical models and uses of the dataset

**Corresponding author: E-mail: jitiprova.ghosh@gmail.com;*

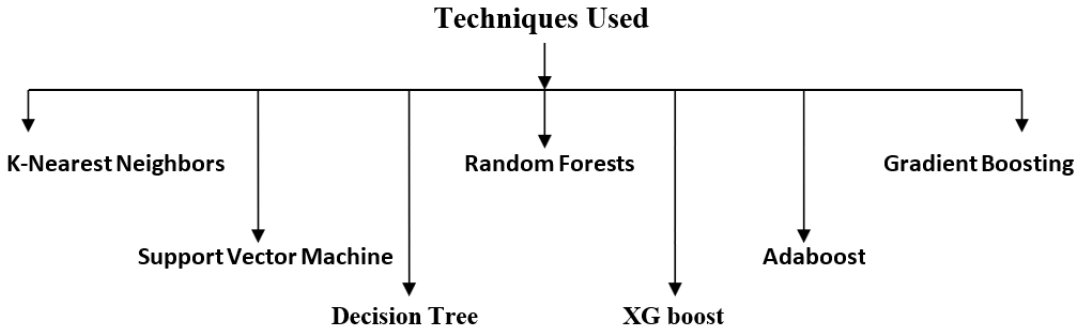
of some previous years having the weather parameters, i.e temperature, dew, humidity, pressure, visibility, wind speed and wind direction also. The terms of forecasting weather in Machine learning methods may be 'No rain', 'Fog', 'Rain' 'Thunderstorms', 'Fog Rain', 'Tropical Cyclones', "Heat waves', 'Cold snaps', 'Heavy downpours', 'Floods' etc. In this research process the predicted outcomes are continuous numeric values temperature [6,7].

The different Machine. Learning Methods which is used in this technique are random forest, decision tree, support vector Machine, KNN, Adaboost, Xgboost, Gradient Boosting, Naïve Bays and logistic regression etc [8-12]. The Models are evaluated in the basis of their accuracy in prediction and verification process. The used techniques of Machine learning method for weather forecasting is analysed in respect of their accuracy level to get an overall idea and its applications [13-15].

2. METHODOLOGY AND TECHNIQUES

Here we aim to use Machine Learning techniques to predict the weather day by day and hourly. We use the techniques namely Chi square test and Naïve Bayes statistics applied on the dataset to examine the useful information from that dataset. First of all here we collect the data by user information. The we process the data by transformation and combination and prepare a database. By Data Mining Techniques the data is stored. At last by making Decision Tree, here we process the dataset and gives the dataset for weather prediction.

The techniques that are used in this machine learning process depicted below in the chart and described briefly to understand the process.



(i) K-nearest neighbors (KNN): KNN method may be used for both regression and classification. Mainly it is utilized for classification. The right class of testing data is being predicted by KNN of a given set of data of various classes by calculating the distance between both the testing data and all training points. After that K point is choosen which are most similar to the test.

(ii) Support Vector Machine: This technique is used for the predictive data processing. Also for the purpose of classificationthe support vector machine is utilized.

(iii) Decision Tree: Among the effective techniques of machine learning, decision tree is the most effective and best used techniques. An item or scenario with a set of attributes is used as an input and provides a 'yes/no' decision in the decision tree method.

(iv) Random Forests: A Collection of decision trees is called as Random forests strategy which generates more tree variety.

(v) XG boost: This method is also known as "Extreme Gradient Boosting." To Solve a number of data science problem more precisely XGboost use parallel tree boosting technique.

(vi) Adaboost: Adaboost a statistical categorization method which is used for the improvement of theresults and is combined with a number of learning algorithms.

(vii) Gradient Boosting: Among the other boosting techniques, Gradient Boosting is the prominent techniquewhich corrects the former mistakes.

3. ANALYSIS OF RESULTS

The different models have been used here to predict the weather data of 21 years (1996-2017). For a particular given region, the data of the accuracy of the different models is summarized in Table 1 and its percentage of accuracy is shown in Fig.1 with graphical representation. Based upon the F1 score of the models used, the performance of all the model is summarized in the Table 2 and represented in the Fig. 2.

Table 1. Dataset used in the experiment

SL	Temp.	Dew	Humidity	Pressure	Visibility	Wind	Event
01	28	24	76	1002	05	11	No Rain
02	29	26	85	1003	05	08	No Rain
03	32	26	78	1004	05	11	No Rain
04	31	26	81	1003	04	13	No Rain
05	31	26	86	1001	04	10	Rain, thunderstorm

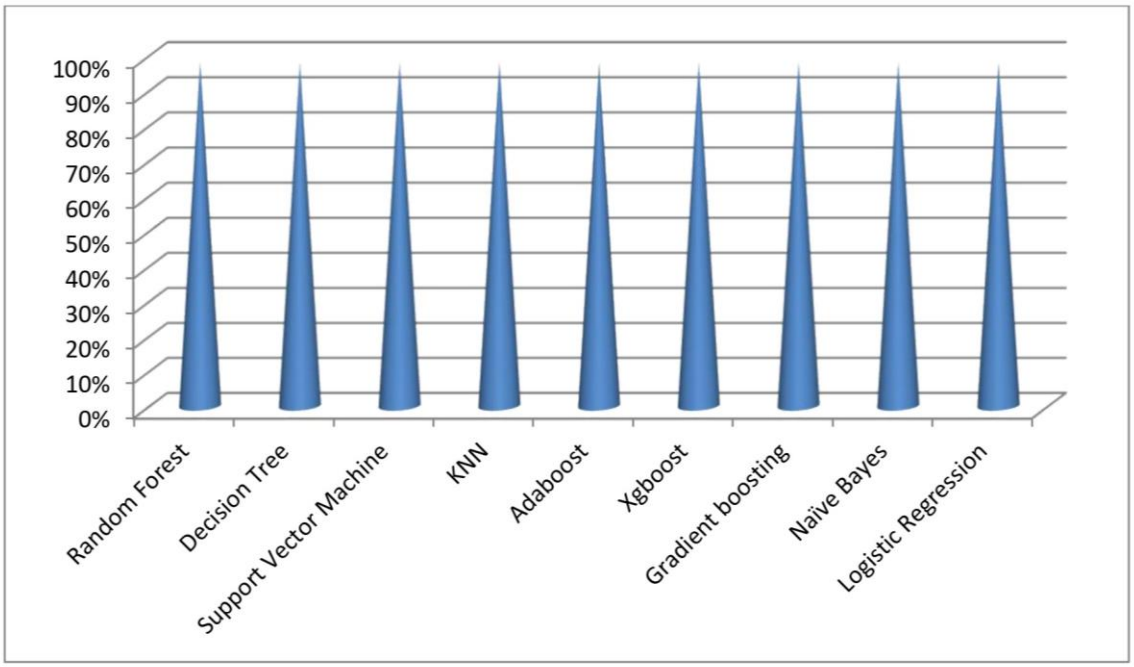


Fig. 1. Graphical comparison of the accuracy of different models

Table 2. F1 Score of different models

Sl No	Model name	F1 score
1	Random forest	0.81
2	Decision Tree	0.75
3	Support Vector Machine	0.74
4	KNN	0.80
5	Adaboost	0.74
6	Xgboost	0.80
7	Gradient boosting	0.83
8	Naïve Bayes	0.79
9	Logistic regression	0.80

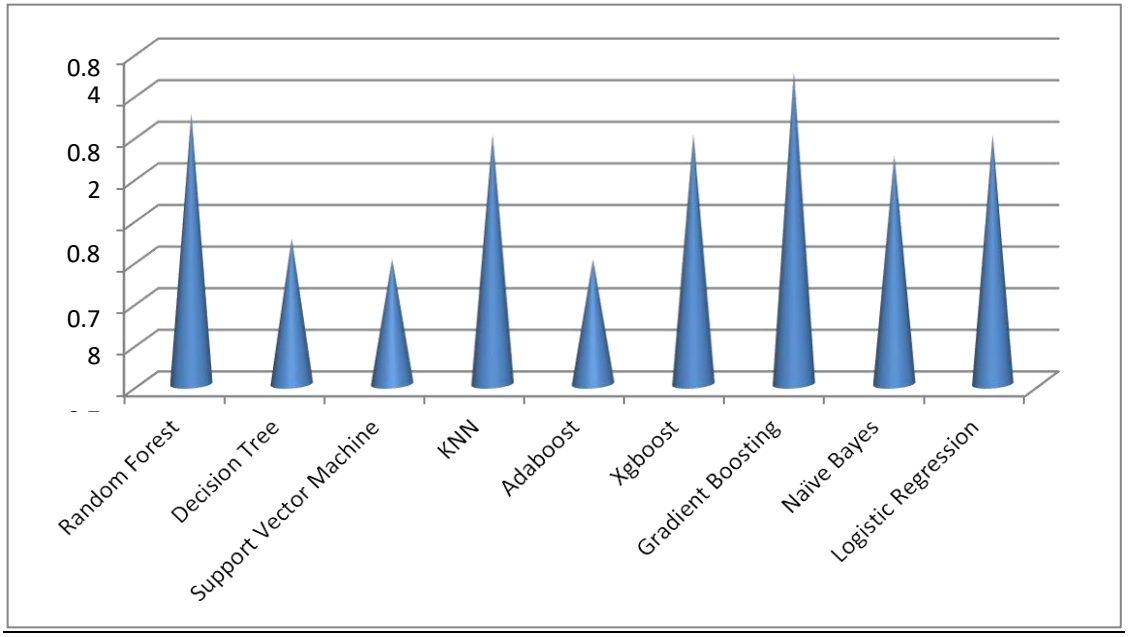


Fig. 2. The graphical comparison of F1 score of different models

A maximum accuracy upto 81.67% observed in these models which needs dataset from predefined resources. Highest accuracy level is observed in Gradient Boosting method. It is evident to use low cost internet of things (IoT) devices in near future such as temperature and humidity sensors to collect the weather data from different parts of a place or an area and to be fed in the machine learning methods.

4. CONCLUSIONS

Machine learning technology can easily be run on almost all the computers including ML devices. Traditional physical models are more resource oriented than machine learning technology. It can provide intelligent models and are much more simpler in comparison to machine learning technique. ML technique can predict weather parameters more precisely and accurately than the traditional models.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Schultz MG, Betancourt C, Gong B, Kleinert F, Langguth M, Leufen LH, Mozaffari A, Stadtler S (2021) Can deep learning heat numerical weather prediction. Phil Trans R Soc A 379:20200097. Available:https://doi.org/10.1098/rsta.2020.0097
2. Sharma A, Agarwal S. Temperature prediction using wavelet neural network. Res J Inf Technol. 2012;4:22-30. Available:https://doi .org/10.3923/rj it.20 I 2.22.30. Available:https://sc ialert.net/abstract/?doi=rj it. 2012.22.30
3. Singh N, Chaturvedi S, Akhter S. Weather forecasting using machine learning algorithm. In: International conference on signal processing and communication (ICSC). 2019;171-174. Available:https://doi.org/10.1109/ICSC45622.2019.8938211
4. Khajure S, Mohod SW. Future weather forecasting using soft computing techniques. Proc Comput Sci. 2016;78:402-407. ISSN 1877-0509. Available:https://doi.org/10.1016/j.procs.2016.02.081
5. Bhardwaj R, Duhoon V. Weather forecasting using soft computing techniques. In: International conference on computing, power and communication technologies (GUCON). 2018:1111-1115. Available:https://doi.org/10.1109/GUCON.2018.8675088
6. Haghbin M, Sharafati A, Motta D. Applications of soft computing models for predicting sea surface temperature: a

comprehensive review and assessment. Prog Earth Planet Sci. 2021;8(4). Available:https://doi.org/10.1186/S40645-020-00400-9

7. Vathsala H, Koolagudi SG (2021) Neuro-fuzzy model for quantified rainfall prediction using data mining and soft computing approaches. IETE...1 Res. Available:https://doi.org/10.1080/03772063.2021.1912648
8. Lin S, Wang G, Zhang S, Li J. Time series prediction based on support vector regression. Inf Technol J. 2006;5:353-357. Available:https://doi.org/10.3923/itj.2006.353.357 Available:https://scialert.net/abstract/?doi=itj.2006.353.357
9. Askari M, Askari H. Time series grey system prediction-based models: gold price forecasting. Trends Appl Sci Res. 2011;6:1287-1292. Available:https://doi.org/10.3923/tasr.2011.1287.1292 Available:https ://sci al ertnet/abstract/?doi=tasr.2011.1287.1292
10. Lee J, Lee J. Constructing efficient regional hazazdous weather prediction models through big data analysis. IJFIS. 2016 16:1-12. Available:https://doi.org/10.5391/15TS.2016.16.1.1
11. Jayasingh SK, Mantri JK, Pradhan S. Weather prediction using hybrid soft computing models. In: Udgata SK, Sethi S, Srirama SN (eds) Intelligent systems. Lecture notes in networks and systems, vol 185. Springer, Singapore; 2021. Available:https://doi.org/10.1007/978-981-33-6081-5_4
12. Sofian IM, Affandi AK, Iskandar I, Apriani Y. Monthly rainfall prediction based on artificial neural networks with backpropagation and radial basis function. Int J Adv Intell Inf. 2018;4(2):154-166. ISSN 2548-3161. Available:https://doi.org/10.26555/ijain.v4i2.208
13. Balogh B, Saint-Martin D, Ribes A. A toy model to investigate stability of AT-based dynamical systems. Geophys Res Lett. 2021;48(8). Available:https://doi.org/10.1029/2020GL092133
14. Jayasingh SK, Gountia D, Samal N, Chinara PK. A novel approach for data classification using neural network. IETE J Res; 2021. Available:https://doi .org110.1080/03772063 .2021.1986152
15. Litta AJ, Idicula SM, Mohanty UC. Artificial neural network model in prediction of meteorological parameters during premonsoon thunderstorms. lnt J Atmos Sci 2013:14. Article ID 525383. Available:https://doi.org/10.1155/2013/525383

Peer-review history:
The peer review history for this paper can be accessed here:
https://www.sdiarticle5.com/review-history/110125